#33
B.O.
3-3-03

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: James CURRY, et al.	
Serial No.: 08/598,457	Group Art Unit: 2644
Filed: February 8, 1996	Examiner: X. Mei
For: SPATIAL SOUND CONFERENCE SYSTEM AND APPARATUS	Atty. Dkt. No.: 414.103/09504869

APPEAL BRIEF

TABLE OF CONTENTS

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: James CURRY, et al.	
Serial No.: 08/598,457	Group Art Unit: 2644
Filed: February 8, 1996	Examiner: X. Mei
For: SPATIAL SOUND CONFERENCE SYSTEM AND APPARATUS	Atty. Dkt. No.: 414.103/09504869

APPEAL BRIEF

BOX AF
Hon. Assistant Commissioner of Patents
Washington, D.C. 20231

Sir:

Appellant hereby submits an original and two copies of this Brief to the Board of Patent Appeals and Interferences in response to the Office Action dated February 8, 2002, from which a Notice of Appeal was filed August 7, 2002. A check is attached covering the fee for filing the Brief in support of the instant appeal together with the extension of time fee. Should any other fees be due, the Commissioner is authorized to withdraw the appropriate fee from Fulbright & Jaworski Deposit Acct. No. 06-2375 under order no. 414.103/09504869. Please date stamp and return the enclosed postcard to evidence receipt of this document.

I. REAL PARTY IN INTEREST

The real party in interest is the assignee, Verizon Services Corp.

II. RELATED APPEALS AND INTERFERENCE

There are no interferences or appeals for related cases.

III. STATUS OF THE CLAIMS

Claims 1-27 were filed with the original application. Claim 24 was amended in appellants' first response in the initial application, of which this is a second Continued Prosecution Application. Thus, claims 1-27 are pending, stand rejected and are appealed. A copy of the appealed claims is attached as APPENDIX 1 to this brief.

IV. STATUS OF THE AMENDMENTS

The sole amendment to the claim, that being to claim 24, was earlier entered in the patent application. There are no outstanding amendments to the claims.

V. SUMMARY OF THE INVENTION

The present invention is directed to a spatial sound conference system that enables participants in a teleconference to distinguish between and identify speakers based on virtual location cues (see page 4, lines 6-9). Spatial sound information may be captured using a dummy head 101 at a conference table 120 (see page 4, lines 9-10; Figures 1, 1A), or spatial sound information may be added to a participant's monaural audio signal using head-related transfer functions (page 6, lines 17 -18) based on an assigned virtual location of a speaker (see page 4, lines 10-13; page 11, line 27 – page12, line 3; Figure 2). Spatial sound signals may be reproduced on spatially disposed loudspeakers preferably positioned near the ears of a listener (see page 4, lines 13 –1 4; page 7, lines 10 –13; Figures 1 and 1A). The spatial sound conference system is designed to enable conferences across a digital network (see page 4, lines 14 –1 6; page 5, line 20; page 6, lines 19 – 28). In addition to audio conferences, the system can provide

spatial sound to audiovisual conferences, long distance learning systems, or virtual reality environments implemented across a network (see page 4, lines 16 –18)

Head-related transfer functions (page 6, lines 17 -18) may be used to simulate the frequency response of audio signals across the head from one ear to the other ear to create a spatial location for a sound (see page 4, lines 19 – 21; page 11, line 27 – page 12, line 3). A computer-generated head-related transfer function (page 6, lines 17 -18) convolved with a single audio signal creates left and right audio signals (page 6, lines 13 - 28) with a spatial sound component (see page 4, lines 21 – 23). Inserting a spatial sound component in a teleconference enhances a listeners' ability to understand a particular speaker (even other than the loudest speaker) including during periods of interruption and overtalk (see page 4, lines 25 – 28).

For example, claim 1 recites:

> 1. A spatial sound conference system comprising:
> a conference station (100) comprising:
> right and left spatially disposed microphones (103, 105, page 5, lines 21 - 22) connected to a communications channel (ISDN facilities 150, page 6, lines 8 - 11, lines 19 - 26; page 12, lines 3 - 6; page 15, lines 6 -22) for receiving right and left audio signals (page 6, lines 13 – 28), wherein the differences between the right and left audio signals represent a head-related transfer function (page 6, lines 17 -18); and
> a remote station (199, page 7, lines 7 - 28) comprising:
> right and left spatially disposed loudspeakers (113, 115; page 7, lines 10 - 18) connected to the communications channel (ISDN facilities 150).

(Note that references to the Figures and specification are illustrative of an embodiment of the invention and are not intended to limit the scope of the claims.)

Independent apparatus claims 15 and 24 and independent method claims 13, 19, 22 and 27 further recite:

> 13. A method for conducting a spatial sound conference comprising the steps of:

converting audio information into right and left audio signals at a conference station (page 6, lines 13 – 17), wherein the conversion imparts a differential characteristic to the right and left audio signals (page 6, lines 17 – 18), and the differential characteristic is represented by a head-related transfer function (page 6, lines 17 - 18; page 4, lines 19 – 25), and the right and left audio signals comprise spatialized audio (page 11, lines 28 – 31; page 12, lines 30 – 31);

transmitting audio information representative of said spatialized audio from the conference station across a communications channel (ISDN facilities 150, page 6, lines 8 -11, lines 19 - 26; page 12, lines 3 - 6; page 15, lines 6 - 22) to a remote station (199, page 7, lines 7-28; page 6, lines 19 - 24); and

playing the spatialized audio in the remote station (199, page 7, lines 7 – 28; page 7, lines 7 – 12).

15. A spatial sound conference system comprising:

a transmitting station (310; page 14, lines 22 - 29) comprising:

a microphone (307; page 14, lines 28 –29) connected to a communications system for receiving an audio signal;

a head-related transfer function unit (335; page 16, lines 7 - 12) connected to the communications system (350, page 15, lines 6 – 7) for imparting a head-related transfer function (page 15, lines 6 - 9) to the audio signal to produce a spatialized audio signal (page 15, lines 13 –23); and

a receiving station (310) comprising:

right and left spatially disposed loudspeakers (303, 305; page 14, lines 24 – 26) connected to the communication system (350) for receiving the spatialized audio signal (page 16, lines 23 – 27).

19. A method for conducting a spatial sound conference comprising the steps of:

receiving an audio signal at a transmitting station (page 14, lines 28 – 29)

transmitting the audio signal from the transmitting station to a spatial sound conference bridge (page 15, lines 6 – 8);

imparting a head-related transfer function (page 6, lines 17 -18) to the audio signal to create a spatialized audio signal (page 15, lines 7 – 9; page 16, lines 3 - 10);

sending the spatialized audio signal from the spatial sound conference bridge to a receiving station (page 16, lines 19 – 22); and

playing the spatialized audio signal on spatially disposed loudspeakers at the receiving station (page 16, lines 25 – 27).

22. A method for conducting a spatial sound conference comprising the steps of:

receiving an audio signal at a transmitting station (page 14, lines 28 – 29);

transmitting the audio signal from the transmitting station to a receiving station (page 16, lines 19 – 22);

imparting a head-related transfer function (page 6, lines 17 -18) to the audio signal to create spatialized audio signal (page 15, lines 7 – 9; page 16, lines 3 – 10; page 16, lines 27 – 30; page 17, lines 4 – 5);

playing the spatialized audio signal on spatially disposed loudspeakers in the receiving station (page 16, lines 25 – 27).

24. A spatial sound conference bridge comprising:

at least two input ports (201, 202, 203, 204) for receiving at least two audio signals (page 12, lines 3 – 6);

a head-related transfer function unit (205) connected to the at least two input ports (201, 202, 203, 204) for imparting a head-related transfer function to at least one received audio signal to produce at least one spatialized audio signal (page 12, lines 25 – 27, lines 30 - 31); and

at least two output ports connected to the head-related transfer function unit for transmitting the spatialized audio signal (page 13, lines 18 -19).

27. A method for conducting a spatial sound conference comprising the steps of:

receiving at least two monaural audio signals (page 12, lines 3 – 6);

generating at least two sets of spatialized audio signals from the at least two monaural audio signals using at least two head-related transfer functions (page 12, line 23 – page 13, line7);

compiling at least one composite signal from the at least two sets of spatialized audio signals (page 13, lines 8 – 18);

transmitting at least one composite signal to a location (page 13, lines 18 – 19); and

playing at least one composite signal at the location. (page 16, lines 23 – 27).

VI. ISSUE ON APPEAL

Whether the Examiner erred in finally rejecting claims 1-27 under 35 U.S.C. § 103(a) as being obvious in light of a combination of Minami et al., in view of Suzuki et al., Cooper, or Hansen, et al..

VII. GROUPING OF THE CLAIMS

The claims stand or fall together with regard to the rejection.

VIII. SUMMARY OF THE ARGUMENT

The Examiner has maintained the rejection of all pending claims under 35 U.S.C. §103(a) as described in the Office Action mailed February 8, 2002 and reaffirmed in the Advisory Action mailed September 10, 2002. The Examiner contends that Minami et al. (U.S. Patent No. 5,555,310, hereinafter, Minami) in view of Suzuki et al. (U.S. Patent No. 5,440, 639, hereinafter, Suzuki), Cooper (U.S. Patent No. 4,910,779, hereinafter, Cooper) or Hansen, et al. (U.S. Patent No. 5,498,478, hereinafter, Hansen, et al.) render all pending claims obvious under 35 U.S.C. §103(a). According to the Examiner, Minami teaches a two-way communication or conference system providing virtual audio or stereo voice transmission to a remote listener, with a pair of microphones located at the conference station and corresponding loudspeakers connected to a communications channel at a the remote station. The Examiner has taken the position that it would have been obvious to modify the teachings of Minami to incorporate the head related features of Suzuki, Cooper of Hansen, et al. "to improve sound localization". (Office Action mailed February 8, 2002, paper # 27, at page 3.)

The rejection must be reversed. Not only is the rationale advanced by the Examiner for making the combination conclusory and without foundation, but contrary to the teachings of the references which teach away from the combination. For example, adding a head-related transfer function to the Minami reference destroys the intended use of the reference, i.e., that of reducing the cost of transmitting a stereo voice signal across a telephone line. Furthermore, Minami describes preventing fluctuation of sound image localization by transmitting a monaural voice transmission when one person is speaking and transmitting a stereo voice transmission when more than one person is speaking, likewise teaching away from imparting the head-related-transfer function/spatial components to the monaural voice signal.

Further, even if the combination asserted by the Examiner were made, it would still fail to render the subject matter of the independent claims obvious, failing to describe or suggest that the differences between the right and left audio signals represent a head-related transfer function. For these reasons, as is more fully presented below, the rejections must be reversed.

IX. ARGUMENT

The rejection of claims 1-27 under 35 U.S.C. § 103(a) as being obvious over Minami et al. in view of Suzuki et al., Cooper or Hansen, et al. is improper.

Claims 1-27 have been rejected under 35 U.S.C. § 103(a) as being obvious over Minami et al. (U.S. Patent No. 5,555,310, hereinafter, Minami) in view of Suzuki et al. (U.S. Patent No. 5,440, 639, hereinafter, Suzuki), Cooper (U.S. Patent No. 4,910,779, hereinafter, Cooper) or Hansen, et al. (U.S. Patent No. 5,498,478, hereinafter, Hansen, et al.). The Examiner asserts that it would have been obvious to combine the stereo voice transmission apparatus of Minami with the head related features or functions of Suzuki, Cooper or Hansen, et al.. According to the Examiner, Minami teaches a two-way communication system or conference system which provides virtual audio or stereo voice transmission for the listener, the Examiner taking the position that it would have been obvious to replace the transfer function "imparted" in Minami with the head related features or functions of Suzuki, Cooper or Hansen, et al..

However, to establish a prima facie case of obviousness there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. No such motivation is present here. For example, claim 1 recites a conference station comprising "right and left spatially disposed microphones connected to a communication channel for receiving right and left audio signals, wherein the differences between the right and left audio signals represent a head-related transfer function". According to claim 1, the differences between the

right and left audio signals represent a head-related transfer function. Minami neither teaches nor suggests such a head-related transfer function. Further, there is no motivation to modify the Minami reference to include such a feature or otherwise to combine Minami and the teachings of the cited secondary references. To the contrary it is clearly improper to combine references where, as here, the references teach away from their combination. In re Grasselli, 713 F.2d 731, 743, 218 USPQ 769, 779 (Fed. Cir. 1983). Adding a head-related transfer function to the Minami reference is contrary to, and even destroys the intended use of, the reference; i.e., to reduce the cost of transmitting a stereo voice signal across a telephone line by selectively transmitting a monaural signal.

Furthermore, Minami describes preventing fluctuation of sound image localization by transmitting a monaural voice transmission when only one person is speaking while transmitting a stereo voice transmission when more than one person is speaking. This method teaches away from imparting the head-related-transfer function/spatial components to the monaural voice signal as to do so would consume further bandwidth.

More specifically, Minami discloses an apparatus where "stereo voice transmission is performed in the multiple simultaneous utterance mode, and monaural voice transmission is performed in a single utterance mode." As pointed out in the background of Minami (column 2, lines 50-58), when a monaural voice transmission is directly applied to a multiple utterance, the sound image localization has an undesirable fluctuation. Minami's stated objective is to prevent fluctuations of the sound image localization, while reducing the cost of transmission, by transmitting the single and multiple utterances by different transmission methods and/or different transmission rates.

Thus, Minami transmits a single utterance as a monaural voice signal and multiple simultaneous utterances as a stereo voice transmission to realize a high band compression ratio. During a conversation when one person is speaking (i.e., single utterance), Minami transmits the speech as a monaural voice signal (column 13, lines 39-65). Minami does not alter this signal with any spatial information before transmission. During a conversation when more than one person is speaking (i.e., multiple utterances), Minami transmits the speech as a stereo voice transmission. However, Minami does not alter this signal with any additional spatial information before transmission; there is no head-related transfer function involved.

The Examiner states that the transfer function "imparted" by Minami may not be disclosed as a head-related-transfer function. However, the transfer function in Minami does not "impart" anything to the transmitted signal, i.e., the transfer function of Minami is a ratio of the input signal and the output signal (column 13, lines 35-65). According to the Minami reference, a single voice signal is transmitted when one person is speaking, and the transfer function is used once the signal is received to reproduce the originally recorded signals for the listener. The transfer function does not "impart" any spatial components on the voice signal. To the contrary, the Minami reference "teaches away" from adding a head-related transfer function. This is because the intended use of Minami, i.e., reducing the cost of transmitting a stereo voice signal across a telephone line, is antithetical to imparting a head-related transfer function; adding a head-related transfer function to the Minami reference is counter productive to reducing bandwidth. That is, imparting a head-related-transfer function to the monaural signal would increase the cost of transmission. Furthermore, according to Minami, when a multiple utterance is detected and a stereo transmission is utilized, the original signal is not altered by a transfer

function to impart spatial components. There are no spatial sound information being applied to the original signal (see column 9, lines 45-59) as required by claim 1.

Minami shows adding sound localization information to a received location. This spatial information is not imparted before transmission. The sound image localization information is obtained from a sound image localization control information generator (column 15, line 8 - column 16 line 21). The sound image localization control information generator generates a plurality of pieces of sound image localization control information Lk including, as information, at least one of delay, phase, and gain differences determined in correspondence with the position of the image displayed on the screen. When the plurality of pieces of sound image localization control information Lk are used, for example, as shown in FIG. 11, sound image localization control is performed as if a voice is produced from the position of speaker's mouth of the image 712 on the screen 710. More specifically, the screen 710 is divided into N x M blocks, and sound image localization is controlled in units of blocks. Even when any one of the delay, phase, and gain differences is used, or a combination of the differences is used, this sound image localization control can be performed. Again, no head-related transfer function is involved

In summary, not only is the combination of references asserted by the Examiner improper but, even if made, the combination of the cited references does not teach or suggest that the differences between the right and left audio signals represent a head-related transfer function. Accordingly, appellant respectfully requests reversal of the outstanding rejection of claim 1 under 35 U.S.C. § 103(a).

Claims 2-12 depend directly or indirectly from claim 1 and, as such, include the limitations thereof. Appellant submits that the differences between what is set forth in claims 2-12 including the limitations of claim 1 from which they depend, are such that they would not

have been obvious to those of ordinary skill in the art in view of the cited references. Reversal of the rejection under 35 U.S.C. § 103(a) of there claim is likewise respectfully requested.

Claim 13 recites a method of converting audio information into right and left audio signals at a conference station, wherein the conversion imparts a differential characteristic to the right and left audio signals, and the differential characteristic is represented by a head-related transfer function. The right and left audio signals comprise spatialized audio and transmitting audio information representative of said spatialized audio from the conference station across a communication channel to a remote station. As described *supra*, adding a head-related transfer function to the Minami is neither described nor suggested by the reference. Reliance on the applied secondary references is also misplaced, there being no motivation to make the combination. To the contrary, the primary reference teaches away from any such a modification again rendering the rejection under 35 U.S.C. §103(a) improper. Further, even if there were motivation to modify Minami as asserted by the Examiner, the claimed invention would not result. For example, claim 13 requires a conversion step that imports a differential characteristic to the right and left audio signals, an feature absent from the asserted references.

Because the combination of the cited references are improper and, even if made, fail to teach or suggest that the differences between the right and left audio signals represent a head-related transfer function, reversal of the outstanding rejection of claim 13 under 35 U.S.C. § 103(a) is appropriate.

Claim 14 depends directly from claim 13 and, as such, includes the limitations thereof. Appellant submits that the differences between what is set forth in claim 14 including the limitations of claim 13 from which it depends, is such that it would not have been obvious to

those of ordinary skill in the art in view of the cited references. Withdrawal of the rejection under 35 U.S.C. § 103(a) is respectfully requested.

Apparatus claim 15 recites a head-related transfer function unit connected to the communications system for imparting a head-related transfer function to the audio signal to produce a spatialized audio signal". As explained in connection with independent apparatus claim 1, the outstanding rejection is both based on an improper modification of the primary Minami reference for which there is no motivation and a combination of references that collectively fail to teach or suggest all of the recited elements of the rejected claim. Accordingly, reversal of the rejection of claim 15 under 35 U.S.C. §103(a) is appropriate and is respectfully solicited.

Claims 16-18 depend directly from claim 15 and, as such, include the limitations thereof. Appellant submits that the differences between what is set forth in claims 16-18 including the limitations of claim 15 from which they depend, are such that they would not have been obvious to those of ordinary skill in the art in view of the cited references. Again, reversal of the outstanding rejection under 35 U.S.C. § 103(a) is respectfully requested.

Method claims 19, 22 and 27 recite subject matter similar to that previously discussed in connection with claim 13 and are considered to be allowable, *inter alia,* for the reasons presented *supra* in connection with claim 13. Specifically, the asserted combination of references is improper under 35 U.S.C. §103(a) and, even if made, would fail to teach or suggest the claimed subject matter of the claims including.

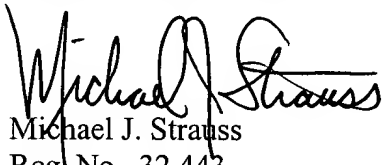
Likewise claims 20 and 21; and claim 23, dependent from claims 19 and 22, respectively, are considered to recite allowable subject matter by, *inter alia*, incorporating the allowable subject matter of these independent claims.

Likewise, independent apparatus claim 24, together with claims 25 and 26 dependent therefrom, are considered to be allowable for the reasons presented *supra* in connection with claim 1.

X. CONCLUSION

Appellant respectfully submits that the rejection of claims 1-27 under 35 U.S.C. §103(a) is not sustainable for the reasons set forth herein. Reversal of the rejection is respectfully solicited.

Respectfully submitted,



Michael J. Strauss
Reg. No. 32,443
Attorney for Appellants

FULBRIGHT & JAWORKSI, LLP
801 Pennsylvania Avenue, N.W.
Washington, D.C. 20004-2615
(202) 662-4632
(512) 418-3000

Date: February 7, 2003